

Sydney Levitus
NODC/NOAA

Project: Ocean Data Archaeology and Rescue

1) Basic Description of Data Set(s):

- A- specification of the specific variables contained in the data set(s) (e.g., sea surface temperature and pressure)**

Vertical ocean profiles of temperature, salinity, oxygen, nutrients, chlorophyll, pH, alkalinity.

- B- the type of observations used in the data set production (e.g., satellite IR, land surface rain gauge)**

Reversing thermometer, samples from water bottles, CTDs, XBTs, MBTs, buoys, profiling floats.

- C- the geographic area covered**

World Ocean

- D- the temporal and spatial resolution of the data set(s)**

Individual profiles at discrete times and locations made by ships and other platforms including buoys and airplanes.

- E- duration of the data set(s)**

Year 1772-present.

- F- standard interval for adding new data (e.g., daily via GTS, delayed mode, monthly update)**

Data are made available as part of the World Ocean Database which is compiled and distributed every 3-4 years.

- G- mechanisms maintained for accessing the data**

Data are available on-line via NODC.WDC website and via CD-ROMs

- H- current uses of data set(s) that support operational designation, e.g., required for formal national and international assessment activities**

Data and products based on the data are widely used by the international scientific research and operational forecasting communities as indicated by citations in scientific journals. Data are used in diagnostics studies of the role of the ocean in climate change, for validating OGCM and AOGCM simulations, as initial, internal, external boundary conditions. Work based on the World Ocean database and products based on WOD have been cited in IPCC and NRC reports. Seasonal-to-interannual forecasting models including the ocean component of the climate system also use these data.

2) Scientific Stewardship Activities Required for Continued Production of the Climate-Quality Data Set

A- quality control procedures, including ongoing improvements

QC procedures are documented by the following two technical reports:

Conkright, M. E., T. P. Boyer, and S. Levitus 1994: Quality control and processing of historical oceanographic nutrient data. NOAA Technical Report NESDIS 79, National Oceanographic Data Center, Wash., D.C., 75 pp.

Boyer, T. P. and S. Levitus, 1994: Quality control of oxygen, temperature and salinity data. NOAA Technical Report No. 81, National Oceanographic Data Center, Wash., D.C., 65 pp.

Many QC problems we encounter are not amenable to identification by statistical procedures as they are problems with incorrect metadata such as bad locations, dates, ship names, etc. This is labor-intensive, time-consuming work that has to be done but which is critical for the production of quality database.

B- bias identification and processing which should clearly explain methods to be used and why undefined analyses or research to develop new bias detection algorithms should be proposed outside of the ARC

Bias identification can be handled if proper metadata such as instrument type and method used to measure samples is available with the data. We put a substantial effort into acquiring such metadata if not available with the data we receive.

C- reprocessing work underway, *e.g.*, refresh rate (version control), production of data set(s) at finer resolution, employment of new processing algorithms, gridding, etc.

Reprocessing including quality control is literally a continuous process. As soon as a database is released users inform us of problems and we find additional problems not previously identified by us. All corrected data are made available on-line at the end of each calendar month.

- D- basic, “hands-on” utilization activities by involved scientists needed to assess the data set(s) quality and initiate prompt, remedial actions if problems are detected**

NODC/OCL scientists process the data and use the data to generate scientific products such as the *World Ocean Atlas* (WOA) climatologies and pentadal to yearly anomaly fields. These scientific analysis activities are critical to producing the highest quality-controlled database we can produce.

- E- identification of data set “point man” or “champion”, i.e., the person whose reputation is most vulnerable with respect to the quality and usefulness of the data set(s).**

Sydney Levitus, NODC/OCL and WDC Director

3) Transition of ARC Project to Operational Center

Outline pathway for eventual transition of your operational process to an established NOAA operational Center using the four steps outlined below.
Steps.

1. operational processing and data archive at PI’s institution only.

NODC is an operational center.

2. data being archived at NOAA Center, but all processing at PI’s institution.

NODC is an operational center.

3. process being run in parallel at PI’s institution and NOAA Center.

NODC is an operational center.

4. processing and archive only at NOAA Center, PI performing Scientific Data Stewardship oversight as needed.

NODC is an operational center.